Artificial Intelligence Technology Empowers Music Creation and Teaching: A Study on the Cultivation Path of Core Competencies for Music Professionals in the New Era

Jie Tian*

Hainan Vocational University of Science and Technology, Haikou, 570203, China
**Corresponding author: nusictjie@163.com*

Abstract: *With the rapid advancement of artificial intelligence technology, its deep integration with the field of music is reshaping the core paradigms of music creation and teaching. This study systematically explores the mechanisms and pathways through which intelligent technologies empower the competency development of music professionals, focusing on their roles in music structure analysis, creative capability reconstruction, and the transformation of teaching systems. By analyzing the applications of music information retrieval, deep learning generative models, and adaptive systems in the musical domain, this research reveals how technology innovates musical grammar and creative models. It further investigates how algorithm-assisted composition, human-computer collaborative creation, and intelligent audio processing expand creative boundaries and reconstruct musical language. On this basis, the study proposes an intelligent teaching framework based on personalized learning path generation, real-time feedback systems, and virtual music mentors to construct a future-oriented music education system. The research demonstrates that artificial intelligence technology not only provides new methodologies for music creation and teaching but also charts a direction for the systematic cultivation of core competencies among music professionals.*

Keywords: *AI music; music information retrieval; human-computer collaborative creation; adaptive learning; virtual music mentors; competency cultivation path*

Introduction

Driven by the wave of intelligent technology, the field of music is undergoing a systematic transformation from creative tools to educational paradigms. The traditional music teaching model and creative methods reveal certain limitations when confronted with new demands for personalization, innovation, and interdisciplinary integration. Meanwhile, the capabilities of artificial intelligence technology—such as data-driven analysis, generative modeling, and interactive adaptability—provide new possibilities for overcoming these constraints. This study, situated at the intersection of musicology, computational science, and educational technology, aims to systematically elucidate the core role of artificial intelligence technology in music structure analysis, creative capability expansion, and the restructuring of teaching systems. It further seeks to explore the cultivation path for core competencies of music professionals empowered by technology. The necessity of this research lies in constructing a theoretical framework for music education adapted to the intelligent era, promoting the transition of music creation and teaching from experience-oriented approaches to a model dually driven by data and cognition, and providing theoretical support and practical guidance for cultivating interdisciplinary music talents equipped with technological literacy and innovative capabilities.

1. Integration Mechanisms of Artificial Intelligence Technology and the Music Field

1.1 Application of Music Information Retrieval Technology in Music Structure Analysis

Music Information Retrieval Technology is dedicated to converting musical audio and symbolic data into computable and analyzable structured information. This technology utilizes feature extraction algorithms to quantitatively model the melodic contours, harmonic progressions, rhythmic patterns, and timbre textures of music. At the level of melodic analysis, fundamental frequency tracking algorithms can discretize continuous audio signals into precise pitch sequences, thereby objectively describing the form and developmental trajectory of musical themes. In the field of harmonic analysis, chord recognition models based on spectral analysis can automatically annotate harmonic functions, revealing the inherent harmonic tension and color variation logic in musical works. Rhythm analysis relies on event detection and beat tracking algorithms, which can reconstruct the temporal structure of music from complex acoustic signals and quantitatively assess its rhythmic stability and complexity. This data-driven structural analysis method surpasses the limitations of traditional auditory analysis, providing high-granularity objective evidence for understanding the organizational principles of musical works, and constitutes a crucial cornerstone for the computationalization of the music knowledge system[1].

1.2 Pattern Construction of Deep Learning Algorithms in Music Generation

Deep learning algorithms utilize multi-layer neural network architectures to perform representation learning on massive music datasets, thereby internalizing the deep grammar and stylistic constraints of music composition. The core of these algorithms lies in constructing generative models capable of capturing temporal dependencies in music. Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, through their internal state transmission mechanisms, effectively model long-term dependencies in the temporal dimension of music, ensuring structural coherence in generated musical phrases. Generative models such as Variational Autoencoders (VAE) map music data to a low-dimensional latent space, enabling the disentanglement and manipulation of high-level semantic features like musical style and emotion. This allows creators to perform semantic interpolation or feature arithmetic within the latent space to explore new musical ideas.

Symbolic music generation models process discrete event sequences, such as MIDI format data, directly learning combinatorial probabilities at the note level. In contrast, end-to-end audio generation models synthesize raw waveforms with rich harmonic structures. These models essentially construct complex probability distributions driven by data, and their generative process represents a systematic behavior of creatively extrapolating and reconstructing learned musical patterns.

1.3 The Interactive Role of Adaptive Systems in Music Cognition Processes

The application of adaptive systems in the field of music focuses on establishing intelligent interactive environments capable of responding to human behavior and dynamically adjusting their feedback. The core mechanism of such systems lies in the closed loop of real-time perception, intent recognition, and strategy optimization. In performance interaction scenarios, the system continuously monitors the performer's musical output and physical expression through audio or motion capture sensors, and infers their performance intent and emotional state in real-time based on pre-trained models[2]. Subsequently, guided by its built-in music knowledge model and interaction strategies, the system generates musically coherent accompanying parts, harmonic fillings, or dynamic effect processing, thereby forming a co-performance relationship.

This interaction is not a static sequence of command and response but rather a process of co-evolution. Through algorithms such as reinforcement learning, the system continuously optimizes its interaction strategies to better align with the performer's style and intent, thereby establishing a profound musical dialogue between human and machine. This dynamic, contextualized interaction provides a new experimental ground for exploring the social and embodied dimensions of music cognition, while also laying a theoretical foundation for constructing highly personalized music learning and creative partners.

2. Intelligent Technology-Driven Reconstruction of Music Creation Capabilities

2.1 Creative Expansion Pathways in Algorithm-Assisted Composition

2.1.1 Music Material Generation Based on Probabilistic Models

Music generation systems based on deep neural networks (such as Transformer models and Generative Adversarial Networks) construct high-dimensional musical representations through latent space learning. These models capture complex stylistic constraints and statistical patterns from training data, generating musical phrases, harmonic progressions, and rhythmic patterns that adhere to musical grammar while demonstrating novelty. By adjusting hyperparameters such as temperature parameters, creators can control the conservativeness or exploratory nature of the generated outcomes, thereby establishing a controllable continuum between stylistic imitation and creative breakthrough.

2.1.2 Creative Control Through Parameterization and Constraint Programming

Parametric composition systems abstract high-level musical features (such as tension, tonal ambiguity, and texture density) into adjustable interface parameters, enabling creators to directly manipulate the macro-form of music rather than individual notes. Constraint satisfaction programming allows creators to define a series of musical rules (such as counterpoint principles or harmonic prohibitions), with the system automatically generating solutions that satisfy all constraints. These two approaches shift the focus of creation from micro-structural design to macro-form shaping and rule system definition, achieving a hierarchical leap in creative logic.

2.1.3 Creativity Optimization Driven by Evolutionary Algorithms

Interactive genetic algorithms place creators in the role of evaluating the system's fitness function. The system first generates a set of musical creative variants, and creators make selections based on subjective aesthetics. The selected variants serve as parent generations to produce new individuals through operations such as crossover and mutation. This iterative selection process enables musical materials to evolve in alignment with the creator's implicit aesthetic preferences, achieving closed-loop collaboration between human intuition and machine computation at the level of creative optimization.

2.2 Evolution of Musical Language in Human-Computer Collaborative Creation Environments

2.2.1 Interactive Paradigms of Real-Time Musical Dialogue

Modern music programming environments and AI collaboration platforms support the establishment of real-time musical dialogues between creators and systems. These systems can instantly interpret, develop, and respond to musical materials input by creators, generating contrapuntal voices, harmonic fillings, or variational developments. This immediate feedback prompts creators to continually re-evaluate and adjust their original musical conceptions, forming a dynamic and emergent creative workflow where the outcomes often exceed the independent expectations of either party[3].

2.2.2 Data-Driven Expansion of Musical Grammar

When analyzing massive music datasets, machine learning models frequently identify implicit patterns and atypical structures that human composers have not explicitly formalized. When these data-driven patterns are integrated into the creative process, they often give rise to musical expressions that transcend the norms of traditional harmony and musical form. Examples include non-standard harmonic progressions, cross-cultural rhythmic fusion, and non-linear formal structures, providing new materials and possibilities for the contemporary development of musical language.

2.2.3 Formation of Aesthetic Decisions with Hybrid Intent

In collaborative creation, the aesthetic qualities of the final musical output are neither purely human intent nor pure machine autonomy. Instead, they emerge as hybrid intentions negotiated through the interactive process. Creators need to develop new critical skills to interpret, evaluate, and integrate musical suggestions provided by the system. This process not only alters the specific form of musical output but also profoundly reshapes the creator's aesthetic judgment and musical decision-making patterns.

2.3 Reshaping Sound Design through Intelligent Audio Processing Technologies

2.3.1 Neural Audio Synthesis and Generation

Deep learning-based neural audio synthesizers (such as DiffWave and WaveGAN) can directly model raw audio waveforms, generating complex sonic textures that cannot be produced by physical models of traditional instruments. Through latent space interpolation and arithmetic operations, these models enable continuous control and exploration of features such as timbre, envelope, and harmonic structures, significantly expanding the sonic foundation for electronic and experimental music.

2.3.2 Intelligent Audio Analysis and Processing

Deep learning-based source separation techniques (such as Demucs and Spleeter) achieve high-precision separation and extraction of individual parts from mixed audio, providing unprecedented flexibility for music restructuring and sampling. Intelligent audio effects perform context-aware processing based on audio content, such as automatically recognizing musical segment structures and applying dynamic effect chains, or mapping spectral characteristics of specific timbres to other sound sources through neural style transfer.

2.3.3 Programmatic Sound Organization and Generation

Artificial intelligence technology enables the programmatic organization of large-scale sound materials. Systems can automatically synthesize, filter, and combine sound segments according to high-level instructions (such as "generate a background soundscape with gradually increasing tension"), constructing complex sonic narrative structures. This semantic-based approach to sound organization liberates sound design from microscopic sample editing, transforming it into a structural element that can directly manipulate macroscopic musical expression.

3. Transformation of Music Education System for the Future

3.1 Intelligent Generation Mechanism for Personalized Learning Paths

3.1.1 Modeling and Analysis of Multi-dimensional Learner Characteristics

The system continuously collects individual performance data, creative outputs, and aural discrimination responses to construct a comprehensive learner profile encompassing skill levels, cognitive styles, error patterns, and aesthetic preferences. Skill dimension analysis includes quantitative metrics such as pitch accuracy, rhythmic stability, and technical fluency; cognitive dimension assessment involves evaluating musical memory capacity, auditory analysis speed, and music reading efficiency; the aesthetic dimension is modeled through the learner's preferences for different musical styles and creative choices. Through feature extraction and cluster analysis, these multi-dimensional data form a holistic understanding of the learner's current state and potential development directions, providing a data foundation for personalized intervention.

3.1.2 Dynamic Knowledge Graphs and Competency Relationship Mapping

Musical competencies are deconstructed into interconnected conceptual nodes and skill units, forming a networked knowledge structure with prerequisite and successive relationships. The nodes in the knowledge graph include not only theoretical knowledge points such as fundamental music theory, harmonics, and musical form analysis but also practical elements like various performance techniques, improvisational skills, and compositional abilities. Through link prediction and community detection algorithms, the graph dynamically reveals the intrinsic relationships and developmental sequences among various competency units. By continuously incorporating group learning data, it optimizes the difficulty gradients and connection weights between knowledge nodes, ensuring the scientific validity and efficiency of competency development pathways[4].

3.1.3 Adaptive Path Planning and Dynamic Adjustment

Based on the Bayesian Knowledge Tracing model, the system continuously evaluates the learner's mastery probability of specific knowledge units and determines the most appropriate challenge level according to Vygotsky's Zone of Proximal Development theory. The path planning engine comprehensively considers the learner's long-term objectives and current state to dynamically select instructional content and practice sequences within the knowledge graph, achieving an optimal balance between learning efficiency and motivation maintenance. The system employs multi-armed bandit algorithms to balance exploration and exploitation, recommending both proven effective learning materials and appropriately introducing novel content to expand competency boundaries, thereby forming a continuously optimized personalized learning trajectory.

3.2 Implementation Models of Real-Time Feedback Systems in Skill Training

3.2.1 Multi-modal Data Acquisition and Feature Extraction

The system integrates computer vision, motion sensing, and audio analysis technologies to comprehensively capture learners' skill performance data. The visual module employs pose estimation algorithms to track the performer's body posture, hand morphology, and movement trajectories, identifying technical flaws and unnecessary muscle tension. Inertial measurement units capture joint angles and movement velocities to quantitatively analyze motion economy and efficiency. The audio processing pipeline performs millisecond-level extraction of features including pitch, rhythm, and timbre, enabling precise diagnosis of intonation deviations, rhythmic instability, and tone quality issues. These multi-source data are synthesized through sensor fusion technology to form a comprehensive representation of skill performance.

3.2.2 Hierarchical Feedback Mechanism and Error Prioritization

The system adopts a hierarchical feedback strategy that distinguishes between core errors and secondary issues, prioritizing the resolution of key technical obstacles affecting musical expression. Utilizing rule engines and machine learning classifiers, the system categorizes detected problems into three severity levels: obstructive errors, developmental deficiencies, and refinement suggestions. Obstructive errors such as severe pitch deviations or chaotic rhythmic structures trigger immediate intervention. Developmental deficiencies including uneven timbre control or limited dynamic variation generate targeted exercises. Refinement suggestions involving subtle aspects of musical interpretation are provided as optional feedback. This hierarchical mechanism ensures learners prioritize solving the most critical technical problems within their limited attentional resources[5].

3.2.3 Musicality Analysis and Expressive Guidance

Transcending basic skill correction, the system incorporates musicality analysis modules to provide data-driven references for artistic interpretation in performance. Through quantitative analysis of parameters including dynamic contours, rhythmic flexibility, and timbral variation, the system can visually present the performer's musical expression characteristics and compare them with reference versions of musical treatment. A natural language generation module converts data analysis results into concrete guidance suggestions, such as "crescendo processes should be more gradual" or "phrase endings require more breath sensation." This data-based musicality feedback cultivates learners' artistic judgment, building a bridge between technical training and musical expression.

3.3 Architectural Design Principles of Virtual Music Mentor Systems

3.3.1 Multi-modal Perception and Contextual Understanding

The perception layer integrates audio analysis, computer vision, and natural language processing modules to achieve comprehensive understanding of learners' musical performance and behavioral intentions. The audio analysis engine employs time-frequency analysis techniques and neural network models to conduct multi-dimensional evaluation of performance quality. The computer vision unit captures body language relevant to performance technique through pose estimation and gesture recognition. The natural language interface supports voice interaction and text comprehension, enabling the system to receive learner questions and feedback. The contextual understanding module synthesizes these multi-modal inputs to construct a unified representation of the current teaching scenario, providing a context-aware foundation for instructional decision-making.

3.3.2 Knowledge Base and Teaching Strategy Engine

The system core comprises a knowledge base encapsulating rich domain knowledge in music and a flexible teaching strategy engine. The knowledge base employs ontological methods to organize semantic relationships among musical concepts, skills, and teaching methodologies, while storing common error patterns and their correction strategies. The teaching strategy engine selects the most appropriate approach from various instructional methods - including explanatory demonstration, guided discovery, and case comparison - based on the learner model and current context. A case-based reasoning mechanism enables the system to retrieve successful cases from historical teaching experiences in similar contexts, providing validated solutions for current problems and ensuring the professionalism and effectiveness of teaching decisions.

3.3.3 Adaptive Interaction and Evolution Mechanisms

The interaction layer employs mixed-initiative dialogue management, both responding to student requests and proactively initiating teaching interventions. The natural language generation module converts internal decisions into coherent dialogues that align with pedagogical logic, supporting various teaching speech acts such as explanation, questioning, and feedback. Through reinforcement learning mechanisms, the system continuously optimizes its teaching strategies, adjusting interaction methods and content difficulty based on student learning outcomes, engagement levels, and long-term development data. This self-evolution capability enables the virtual mentor to gradually adapt to specific learners' changing needs, forming highly personalized teaching styles and ultimately developing into an intelligent teaching partner capable of facilitating deep learning.

Conclusion

This study systematically elaborates on the empowerment mechanisms and practical pathways of artificial intelligence technology in the fields of music creation and teaching, forming a complete theoretical framework for technology empowerment that spans from music structure analysis and creative capability reconstruction to the transformation of teaching systems. Research demonstrates that technology-based methods rooted in music information retrieval and deep learning enable the quantitative analysis and generative expansion of musical grammar and style. Human-computer collaborative creation environments and intelligent audio processing technologies significantly expand the possibilities of musical expression and the boundaries of sound design. Meanwhile, personalized learning path generation, real-time feedback systems, and virtual music mentor systems collectively construct an adaptive music teaching architecture oriented toward the future. These findings not only confirm the systematic support role of artificial intelligence technology in cultivating the competencies of music professionals but also highlight the inevitable trend of deep integration between technology and art. Future research should further focus on establishing cross-disciplinary collaboration mechanisms, exploring the application of explainable AI in creative music processes, and striving to build intelligent music systems that better align with human cognitive and aesthetic development, thereby continuously promoting innovation and evolution in music education and creative practice within the intelligent era.

Fund Projects

Funding "Project supported by the Education Department of Hainan Province, project number: Hnjg2025ZD-71".

References

[1] Luo Xuan. "Research on the Construction of a Quality Assurance System for Training Musicology Majors in Universities in the Digital Intelligence Era." Journal of Social Science of Jiamusi University 43.10 (2025): 150-152.

[2] Feng Jianqin. "Exploration of Applied Talent Training Models for Music Majors in Higher Vocational Colleges Based on Deep Industry-Education Integration." Proceedings of the 2025 Forum on Higher Education Development (Volume I). Ed. Jinken Institute of Technology, 2025, 36-37.

[3] Wang Huijuan. "Research on the Optimization of the Talent Training Structure for Music Education Majors in the New Era." Music Life 05 (2025): 83-87.

[4] Zhang Man. "The Path of Talent Training and Reform for Music Majors in Universities." Popular Art and Literature 17 (2023): 126-128.

[5] Yi Jian. "Research and Practice on the Training Model for Musicology Majors in Higher Education Institutions in the New Era—A Review of Music Talent Cultivation in the Innovative Era." Science and Technology Management Research 43.13 (2023): 250.